PROPERTY OF
CETON UNIVERSITY LIBRARY
RECEIVED APR 19 1941

SW

The
PSYCHOLOGICAL
RECORD . . . .

JANUARY, 1941
Vol. IV No. 14

A FACTOR ANALYSIS OF VERBAL ABILITY

DONALD M. JOHNSON and FLOYD REYNOLDS

THE PRINCIPIA PRESS, INC.
BLOOMINGTON, INDIANA

Price of this number, 35 cents

EDITOR	J. R. KANTOR, *Indiana University*
BUSINESS EDITOR	C. M. LOUTTIT, *Indiana University*

DEPARTMENTAL EDITORS:

ABNORMAL	EDMUND S. CONKLIN, *Indiana University*
CHILD	MARTIN REYMERT, *Mooseheart Laboratory*
CLINICAL	G. A. KELLY, *Fort Hays K. S. C.*
COMPARATIVE	E. A. CULLER, *University of Rochester*
EDUCATIONAL	J. G. PEATMAN, *City College of New York*
EXPERIMENTAL	B. F. SKINNER, *University of Minnesota*
INDUSTRIAL	JOSEPH TIFFIN, *Purdue University*
PHYSIOLOGICAL	C. F. SCOFIELD, *University of Buffalo*
PSYCHOMETRICS	J. P. GUILFORD, *University of Nebraska*
SOCIAL	NORMAN C. MEIER, *University of Iowa*
ASSISTANT EDITOR	J. W. CARTER, JR., *Wichita, Kansas*

The Principia Press, Inc., has undertaken the publication of this co-operative journal to afford authors the opportunity of immediate publication at the least possible expense. The present low cost of publication and possible future reductions, depend entirely upon the number of subscriptions. The subscription price has been made low to induce individuals to subscribe. Under the Articles of Incorporation of the Principia Press no profit can be made on any of its publications. Therefore an increase in the number of subscribers will be reflected in reduced costs to authors and in increase in the number of pages published annually.

EDITORIAL BOARD. The above named board of associate editors have agreed to serve in an active capacity for a period of three years, and all manuscripts will be submitted to them.

MANUSCRIPTS may be sent to the appropriate associate editor or to Dr. J. R. Kantor. Longer papers (30 or more typewritten pages) in the above mentioned fields are especially desired.

COSTS. The cost of publication will be $2.00 per page of approximately 470 words. The charge for line cut illustrations will be $2.00 regardless of size. Special charges will have to be made for complex tabular matter and for half-tone or colored illustrations.

REPRINTS. One hundred copies of each paper will be supplied gratis. Additional copies may be purchased in any quantity at a rate which will be set when the author is informed of the cost of publication.

ROYALTIES. Fifty per cent of the net income from the sale of individual copies, or from copies sold as part of back volumes, will be credited as royalties to the author's account. Royalties cannot be paid on income from subscriptions current in the year of publication.

SUBSCRIPTIONS. The subscription price is $4.00 per year for an annual volume of approximately 500 pages. Individual papers will be sold at an advertised price, which will depend upon the size of the article. Foreign subscriptions will be $4.50.

CORRESPONDENCE concerning editorial matters should be addressed to Dr. J. R. Kantor, Indiana University. Business correspondence should be addressed to Dr. C. M. Louttit, Principia Press, Inc., Bloomington, Ind.

A FACTOR ANALYSIS OF VERBAL ABILITY

By Donald M. Johnson and Floyd Reynolds
Fort Hays Kansas State College

The investigation reported here was undertaken because of the writers' interest in the processes underlying problem-solving. Familiarity with the experimental literature on the higher mental processes, such as Woodworth has summarized in the last two chapters of his recent work (16), suggests that there may be two fundamental processes involved in solving all problems: (1) calling up various acts or responses, and (2) selecting those responses which meet the requirements of the problem.

Comparative psychology teaches that in the first process, calling up responses, the higher, more intelligent species show more variability. They have a larger repertoire of responses. Munn states, for example, that "stereotypy decreases and general plasticity of response increases as the higher forms are approached" (7, p. 100). Clinical psychology shows that the bright use more different methods of performing than the dull. Experiments on human adults give similar results and suggest that individual differences in this process may be an important factor in solving problems. Let us call this hypothetical factor F, for flow of responses.

But problems are not solved by aimless activity alone. The problem-solving experiments, both animal and human, show restriction of the responses, under the influence of some sort of directive tendency, to the more likely places, objects or ideas. Complete solution, including transfer, implies recognition of the fit of certain responses to the requirements of the problem. Let us assume that individual differences in this process may be another important factor in solving problems. Let us call this hypothetical factor S, for selection of responses.

It would be difficult to set up an experiment to isolate either of these processes. In operation each is complicated by the other. Individual differences in the two processes may be studied, however, by factor analysis. It seemed likely that, if a suitable battery of tests could be given, factor analysis of the scores would yield useful information as to the presence and importance of these processes.

We do not know, of course, whether flow of one kind of

responses, *e.g.* verbal, is related to flow of another kind, *e.g.* numerical. Hence, we have confined our investigation to verbal tasks. In statistical analyses of intellectual ability the verbal factor has consistently appeared as a factor of considerable importance. (See 1, 2, 4, 6 and 13.) Any information regarding the processes involved in solving verbal problems would, therefore, advance our knowledge of intellectual ability. Our hypothesis is that factor analysis of a number of verbal tests will disclose Factor F, flow of responses, and Factor S, selection of responses, in significant amounts.

The Tests

The selection of tests for factor analysis is an important matter. For present purposes the range of verbal ability should be covered. There should be some tests which could be presumed to include F and S in varying amounts. Yet some of the tests should be so simple that the resulting statistical factors—whatever they might turn out to be—could be interpreted without ambiguity. Ten tests were used, five of which were standard tests, and five constructed *ad hoc*.

1. Flow. This test was intended to measure F, or flow of responses, in as pure a form as possible. It was suggested by a test on the ten-year level of the old Stanford-Binet (11). The directions were to write as many words as possible in two minutes.

2. Multiple Completion. This test was similar to Flow but required some selection or fitting of the responses. Sample:

The horse pulled back ....

Each blank was to be completed by a word which, if used alone, would complete the sense of the sentence. Twelve minutes.

3. Paired Opposites. Instructions were given to name as many words and their opposites as possible in four minutes. This test was expected to depend largely upon flow of responses. Some degree of restriction is evident, however.

4. Word Cross-outs. The subjects crossed out irrelevant words that had been inserted in the text of an article taken from a popular magazine. For example:

Midget auto racing has developed servile a handicapping system based on the . . .

Servile is to crossed out. The content of this test could not be estimated in advance. It was included for the sake of variety. Three minutes.

5. Same and Opposite. This test was taken from the 1931 A. C. E. examination (15). There are four words on a line, two of which are either the same or opposite in meaning. The subjects were to identify the two words. Five minutes. Sample:

1. obtuse, 2. aromatic, 3. acute, 4. occult.

6. Verbal Relations. Two types of verbal problem were used in this test. Samples:

(a) Plank ..... meals
(b) Irksome is to wearisome as engrossed is to .....
1. enhanced, 2. occupied, 3. apathy, 4. dilapidated.

The problem in (a) was to find a word similar in meaning to both plank and meals, *e.g.* board. This type of item was suggested in a recent report by Shipley (10). The (b) items were of the usual analogy type. Fifteen minutes.

7. Vocabulary. The vocabulary section of the Nelson-Denny Reading Test (8) was used to get vocabulary scores. It contains five-choice items, such as:

To feign is to 1. fret, 2. faint, 3. molest, 4. pretend, 5. portend.

Ten minutes.

8. Restricted Completion. This test was the completion test of the 1931 A. C. E. examination (15). Sample:

A dead body, especially a human one, is sometimes called a (7) .....

The number in parenthesis indicates the number of letters in the most appropriate word. It is here called a "restricted completion" test because the response is restricted in two ways: meaning and length. In Test 2, Multiple Completion, there is much less restriction or fitting. Eight minutes.

Factor analysis of these eight tests could be expected to disclose some information about individual differences in verbal ability. If a small number of factors should be found, it would be interesting to investigate their involvement in a practically important verbal skill, such as reading. Therefore a reading test was included. An intelligence test was included because the items on intelligence tests are usually minor problems to be solved, and because intelligence-test scores are usually closely related to verbal proficiency.

9. Intelligence. The Henmon-Nelson Test of Mental Ability, Form A, for college students was used (5). This is a well-known multiple-choice test. Approximately half of the items seem to be verbal. Thirty minutes.

10. Reading. The Comprehension section of the Nelson-Denny Reading Test, Form A, was used (8). The subjects read paragraphs and answer multiple-choice questions about them. Twenty minutes.

These tests were all given with time limits. The times, except for the standard tests, were selected after a little preliminary experimentation. When the reliability coefficients are presented, it will be evident that some of the tests should have been longer.

Subjects

The tests were given during three class hours to 120 summer-school students at Fort Hays Kansas State College. The group was somewhat heterogeneous in respect to age since some of the subjects were teachers who had been teaching a few years. The records of 113 subjects were complete and constitute the data for the factor analysis.

Statistical Treatment

The ten tests were scored, and means and SD's for the ten distributions were computed in the usual way. These are presented in Table I. Table I also shows the reliabilities of each test. For the standard tests the reliability reported by the author was used. In the case of the Nelson-Denny Reading Test (8), which consists of two parts, Vocabulary and Reading, the reliability of the total score only

TABLE I

MEANS, DISPERSIONS AND RELIABILITIES FOR 113 COLLEGE STUDENTS ON TEN VERBAL TESTS

	Test	Mean	SD	Reliability Coefficient
1.	Flow	54	8.3	.603
2.	Multiple Completion	137	34.4	.885
3.	Paired Opposites	21	4.8	.405
4.	Word Cross-outs	21	7.4	.844
5.	Same and Opposite	11	6.2	.87
6.	Verbal Relations	16	4.3	.785
7.	Vocabulary	43	15.2	.842
8.	Restricted Completion	18	6.5	.81
9.	Intelligence	44	10.6	.89
10.	Reading	20	6.3	.842

is given. The assumption was made that both parts are of approximately equal reliability, and the Spearman-Brown "prophecy" formula was used to estimate the reliability of each part. For the newly constructed tests reliability was computed by the split-half method with the aid of the Spearman-Brown formula. Since two kinds of items were used in Test 6, items 1, 2, 5, 6, 9, 10 . . . were correlated against items 3, 4, 7, 8, 11, 12 . . . . In Tests 1, 2, 3 and 4, where the score was simply the number of correct words, the subjects were stopped at half-time and asked to draw a line. Then all were started again at the same place. First and last halves were correlated to get reliability coefficients. Some of these reliability coefficients are probably not highly accurate but they are ac-

TABLE II

CORRELATION MATRIX FOR TEN VERBAL TESTS

Raw Correlations Below the Diagonal; Corrected Above

	1	2	3	4	5	6	7	8	9	10
1		.415	.573	.310	.297	.275	.181	.030	.135	.185
2	.303		.299	.211	.249	.193	.203	.123	.308	.403
3	.283	.176		.387	.517	.567	.411	.469	.153	.238
4	.221	.179	.226		.525	.650	.577	.452	.750	.695
5	.215	.215	.307	.450		.811	.829	.659	.730	.686
6	.189	.158	.320	.529	.670		.688	.754	.821	.642
7	.129	.171	.240	.486	.710	.559		.548	.665	.717
8	.021	.102	.269	.374	.553	.601	.453		.562	.505
9	.099	.269	.092	.650	.642	.686	.576	.477		.764
10	.132	.340	.139	.586	.587	.603	.604	.417	.662	

curate enough for present purposes. Test 3 has very low reliability and does not contribute much to the present investigation. The time for Test 3 was only 4 minutes.

The scores were next transmuted into sigma scores and the intercorrelations among the ten tests computed. Raw and corrected correlations are presented in Table II. Some of the correlations in Table II are rather high and some are almost zero. None is negative.

Thurstone's centroid method of factor analysis (12) was used in the present study. Its convenience and utility have been pointed out by Garrett (2) and by Guilford (4) among others. Guilford's outline (3) was followed. The analysis was carried through with raw r's and with corrected r's. Since the variance due to unreliability varied considerably from test to test, it seemed to the writers that the factor loadings resulting from the use of corrected r's could be more clearly interpreted. Roff (9) has shown that both correlation matrices have the same rank. The analysis was carried through twice. At the close of the second analysis the communalities were quite similar to the communalities used in computing the first factor loadings.

After two factors had been extracted, the distribution of residual correlations had a PE of .056. The PE of a zero r when N = 113 would be .064. We are not justified, therefore, in extracting a third factor, especially since only ten tests were used. The first factor pattern is shown in Table III and plotted in Figure 1.

TABLE III
FIRST FACTOR PATTERN

Test	Factor weights I	II	h^2
1	.303	—.476	.319
2	.361	—.355	.256
3	.380	—.173	.174
4	.697	—.080	.492
5	.810	.232	.710
6	.801	.202	.682
7	.743	.265	.622
8	.619	.281	.462
9	.775	.183	.634
10	.758	—.161	.601

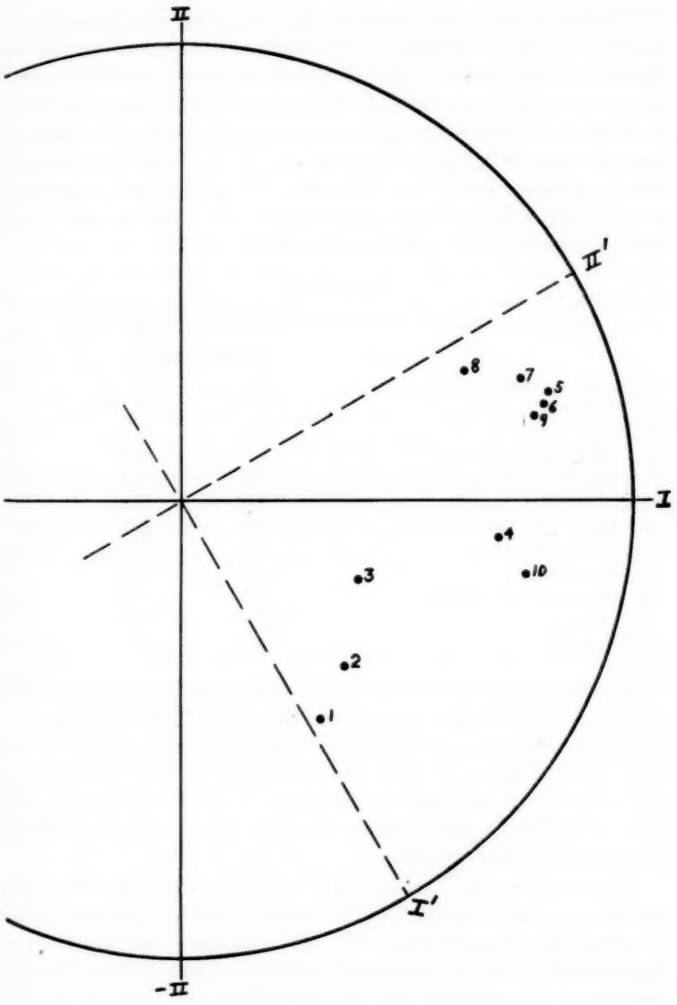

FIGURE 1

Plot of ten test vectors in reference to the centroid axes I and II. After rotation the axes are called I′ and II′.

In rotation only one principle was followed: the principle of positive manifold. Rotation 60° clockwise suffices to render all factor loadings positive. It can be seen from Figure 1 that some leeway is permitted in rotating. Rotation of a little more than 60°, or a little less, would have given a positive manifold. We have no grounds for assuming that either Test 1 or Test 8 measures a factor in purer form than the other. It seemed best, therefore, to place the reference axes approximately the same distance from each. The axes after rotation are called I′ and II′.

TABLE IV
FINAL FACTOR PATTERN AFTER ROTATION

	Test	Factor weights I′ (flow)	II′ (selection)	h^2
1.	Flow	.564	.024	.319
2.	Multiple Completion	.488	.135	.256
3.	Paired Opposites	.340	.246	.174
4.	Word Cross-outs	.418	.564	.492
5.	Same and Opposite	.204	.818	.710
6.	Verbal Relations	.226	.795	.682
7.	Vocabulary	.143	.776	.622
8.	Restricted Completion	.067	.677	.462
9.	Intelligence	.230	.763	.634
10.	Reading	.516	.575	.601

In interpreting the rotated factor pattern one usually looks for similarities of material among the tests with high loadings in a factor. In the present problem the material is presumably of one kind and we therefore look for similarities in ways of handling the material, *i.e.* processes or functions. Test 9, Intelligence, and Test 10, Reading, are relatively complex and are not immediately helpful in the interpretation.

If we look at the loadings in I′ in the light of the hypothesis outlined in our introduction, it becomes easy to identify Factor I′ as our hypothetical Factor F. The tests with the largest loadings in this factor *are* those which are most dependent on flow of responses. Going from Test 1 to Test 8 we see that, as the free flow of responses is restricted, the factor loadings approach zero.

By similar reasoning Factor II′ can be identified as our hypothetical Factor S. The factor loadings *do* increase as one goes

from the tests involving free flow of words to those involving the selection of the words which fit the requirements of the problem. The interpretation of the factor pattern in accordance with our hypothesis seems quite clear-cut.

A secondary purpose of the investigation was to determine to what extent Factors F and S are involved in general intelligence and in reading. Table IV shows that the Henmon-Nelson intelligence test correlates .230 with Factor F, flow of responses, and .763 with Factor S, selection of responses. One would expect S to be a greater factor in intelligence tests than F. It is an interesting finding, nevertheless, in view of the fact that the test is a timed test. This test has fairly high correlations with other group intelligence tests and may be taken as typical.

Factors F and S seem to be about equally involved in the reading scores. The reading test yields one of the highest loadings in Factor F. It is well known that scores on tests of intelligence, vocabulary and reading are rather closely related. The correlations in Table II demonstrate these relationships again. Table IV, however, gives new information about similarities and differences in the processes operating in these three test situations.

Discussion

This is not the place for a critical discussion of factor analysis. Fortunately, several of the controversial issues in factor analysis are evaded by the simplicity of the present experimental design. For example, the rotation which is necessary to clarify the factor pattern is obvious from inspection of Figure 1. There would probably be no difference of opinion on this point. As to the independence of the factors there would be little dispute since Tests 1 and 8, which best define the two factors, show a corrected correlation of only .030. (See Table II.)

It should be noted that some of the communalities are not very large. Extracting two factors leaves much of the variance unaccounted for. Speed of writing might account for some of the variance in the present analysis. A more complete factor analysis, determining the factorial composition of verbal tests in more detail, would use a larger number of tests—and might disclose other important processes. Tests of verbal association and of attention

span for words and concepts should be included. Our purpose in the present investigation was not to do a complete factor analysis, but to determine the importance of two processes which had been postulated by experimental rather than factorial methods. In spite of the small scale of the present study and the inexactnesses which have been pointed out above, the analysis indicates that the two processes considered here, calling up words, and selecting those which meet the requirements of the problem, are involved in solving problems of a verbal nature. The effects of individual differences in each of these processes on the scores in each test is indicated by the factor weights, but the writers would not claim much accuracy for the factor weights because of their dependence upon the selection of tests, selection of subjects, manner of scoring, et cetera. The relative position of the tests in terms of their loadings with each factor is probably fairly accurate.

Materials versus processes. Statistical analyses of intellectual ability have emphasized materials rather than processes or operations. (See 1, pp. 313-315, and 4, pp. 375-382.) This may not be true in the future. It may be that similarities in materials are easier to see than similarities in processes, and thus attract attention first. When the same material is used in all tests, as in the present study, processes are, of course, emphasized. If future factor analyses disclose other processes as important factors in problem-solving, one would expect to hear psychological terms, such as "grasp" and "elaboration of alternatives," rather than logical terms, such as "generalization," "analogy" and "deduction."

An interesting task for the future would be to determine whether the same or similar processes are involved in solving other kinds of problem, *e.g.* numerical. *A priori,* one might expect some similarity in fitting responses to a verbal problem and to a numerical problem, but not in flow of verbal and of numerical responses. Probably S is a more general factor than F.

Note that Tests 5, Same and Opposite, 6, Verbal Relations, and 9, Intelligence, in which the items are minor problems and the responses are highly restricted both formally and functionally, have factor weights in F which appear to be significant. This suggests that flow of relations, concepts, schemas (16, p. 798), hypotheses, insights, plans of action, et cetera, may be included in Factor F.

These are probably all verbal for most people. One biological process could conceivably yield the most logically diverse intellectual products.

Is verbal ability unitary? Tests 5, 6, 7, 8 and 9 are of the kind commonly used in factor analysis. Figure 1 shows that verbal ability is practically unitary as far as these five tests are concerned. The question is whether the other five tests can be ignored in a general description of mental ability. Test 10 is a reading test, with approximately equal weights in both factors. Reading is, of course, an important verbal skill; it cannot be ignored. The other four tests are novel tests, constructed to test a particular hypothesis. But they are certainly verbal tests. It is the writers' opinion that, if verbal ability is considered as the capacity to use verbal material in the solution of verbal problems, at least two dimensions are necessary for its factorial description.

Are other interpretations possible? Since the factor pattern was interpreted and the factors named in the light of a particular hypothesis, it is desirable to look for alternative interpretations. Reconsideration of Table IV suggests that Factor I′ could be called a speed factor and Factor II′ a precision factor. But these are probably merely alternative names for the factors we have previously called by more descriptive names. I′ cannot be an unqualified speed factor because all tests were timed. It must be speed of a sort which is especially needed on Tests 1, 2, 4 and 10.

Professor Thurstone (14) has recently discussed two verbal factors discovered by centroid analysis. V is the well-known verbal factor, and W, a word-fluency factor, "is revealed by readiness in producing words to fit a specified set" (p. 200). The similarity of Thurstone's description of W and our description of F is striking. It is possible that our S is the usual V or verbal factor while F is the newer W or word-fluency factor. If they are the same, their independent discovery is an argument for the validity of the method.

Thurstone adds that "the recognition form of an opposites test, for example, should require little or no W, whereas a free-writing form of the same test should have high saturation in W" (p. 201). This statement holds only superficially for our F. Test 8, which has the least F, is formally a free-writing completion test. Psy-

chologically, however, the flow of responses is highly restricted. Here is an opportunity for confusion between formal and functional characteristics of a task. V and W are significantly correlated (p. 201) while F and S are not. If a wider variety of tests had been included in the present study, would F and S have been correlated?

A methodological note should be added. This investigation was planned as a test of a particular hypothesis. Then the factors were named in the light of this hypothesis. Such a procedure introduces the possibility of bias in selecting the tests and in interpreting the factor pattern, a possibility which is greater with factor analysis than with more objective statistical tools. Nevertheless, the benefits of a hypothesis in planning the investigation and selecting the tests far outweigh the dangers of misinterpretation. Thurstone (14) points out the advantages of developing a hypothesis from a factor analysis, then testing this hypothesis by a subsequent factor analysis. The writers would add that the use of factorial methods to test a hypothesis developed by different methods—in the present case, by experimental methods—has its advantages also.

Summary

The literature on problem-solving among animal and human subjects suggests that there may be two fundamental processes involved in solving all problems: (F) the flow of various acts or responses, and (S) the selection of these responses according to the requirements of the problem. This leads to the hypothesis that individual differences in these processes would be important in determining scores on tests which include the solution of problems.

The investigation was limited to the solution of verbal problems. Factor analysis (centroid method) of ten verbal tests disclosed two independent factors which on rotation can be identified as the hypothetical Factors F and S. The importance of these factors in such common tests as vocabulary, intelligence and reading is shown by the factor weights. It seems quite clear-cut that the two processes or functions suggested by the experimental literature are important in solving verbal problems, and that individual differences in F, flow of responses, and S, selection of responses, are responsible for a large share of the variance in scores on verbal tests. Relations

between the findings of this investigation and problem-solving in general are discussed.

Fort Hays Kansas State College
Hays, Kansas

BIBLIOGRAPHY

1. Anastasi, Anne. Differential Psychology. Macmillan, 1937.
2. Garrett, H. E. Differentiable mental traits. *Psychol. Record* 1938, 2, 259-298.
3. Guilford, J. P. Psychometric Methods. McGraw-Hill, 1936.
4. ..... Human abilities, *Psychol. Rev.* 1940, 47, 367-394.
5. Henmon, V. A. C., and Nelson, M. J. The Henmon-Nelson Tests of Mental Ability. Houghton Mifflin, 1929.
6. Kelley, T. L. Essential Traits of Mental Life. Harvard Univ. Press, 1935.
7. Munn, N. L. Psychological Development, Houghton Mifflin, 1938.
8. Nelson, M. J., and Denny, E. C. The Nelson-Denny Reading Test. Houghton Mifflin, 1929.
9. Roff, Merrill. The relation between results obtainable with raw and corrected correlation coefficients in multiple factor analysis. *Psychometrika,* 1937, 2, 35-39.
10. Shipley, W. C. A self-administering scale for measuring intellectual impairment and deterioration. *Jour. of Psychol.* 1940, 9, 371-377.
11. Terman, L. M. The Measurement of Intelligence. Houghton Mifflin, 1916.
12. Thurstone, L. L. Vectors of Mind. Univ. Chicago Press, 1935.
13. ..... Primary mental abilities. *Psychometric Mono.* 1938, No. 1, Pp. 121.
14. ..... Current issues in factor analysis. *Psychol. Bull.* 1940, 37, 189-236.
15. ..... and Thurstone, T. G. Psychological Examinations. American Council on Education, Washington, 1931 ed.
16. Woodworth, R. S. Experimental Psychology. Holt, 1938.